

Theory and Behavior Analysis

John W. Donahoe

University of Massachusetts, Amherst

I begin my commentary on the contributions of Peter Killeen (2013) and Jack McDowell (2013) by endorsing the general conclusion expressed by Linda Hayes at the end of the Theory and Philosophy conference: These papers are of a very high order and put a lie to the claim that theory plays little role in behavior analysis. My primary purpose here is to discuss the cellular automaton (CA) and genetic algorithms (GA) models described by McDowell in the light of (a) some of the points identified in Killeen's discussion of theory, and (b) my own views on the merits of CA and GA models vis-à-vis neural network (NN) models of behavior.

The three types of models are alike in that each seeks to provide an account of complex behavior as the emergent product of the repeated action of relatively simple processes. In other respects, the models are quite different. The CA model is the product of formal (i.e., logico-mathematical) considerations alone, whereas the GA and NN models are conceptually related to biology; genetics in the first instance and neuroscience in the second. Because both the GA and NN models are inspired by biology, they exemplify Darwinian (selectionist) thinking (Donahoe, 2003). That is, the environment favors or disfavors (i.e., selects) some entities over others. The entities in the GA model are simulated genes within a population of different organisms. The entities in the NN

model are simulated synaptic connections within a population of different behaviors of a single organism. Before moving to a consideration of these specific models in terms of some of Killeen's comments about theory, let me state at the outset that the CA, GA, and NN models are not the only viable approaches to understanding complex behavior. The interpretation of behavior can fruitfully be conducted at various levels of analysis, a matter to which I return at the conclusion of the commentary.

THE EPISTEMOLOGICAL LAYER CAKE

The Empirical Layer

The GA model has as its output variable the mean relative rate of responding observed in a variety of concurrent schedules. As such, the targets of the model are so-called *molar* measures of behavior, that is, the frequency of responding over an experimental session in a group of organisms. Molar measures of responding on concurrent schedules have been the object of theorizing primarily by those influenced by the work of William Baum (e.g., Baum, 1973). In Baum's view, behavior theory "require(s) a simple, but fundamental, change in the law of effect: from a law based on contiguity of events to a law based on correlation between events" (1973, p. 137). Correlational measures require the observation of multiple responses over temporally extended periods.

The molar view contrasts with B. F. Skinner's so-called *molecular* view, in which individual responses are affected by the particular events with which they are contiguous (Ferster & Skinner, 1957; Morse, 1966). In the

Please address correspondence to the author at the Division of Behavioral Neuroscience Department of Psychology, University of Massachusetts–Amherst, Amherst, Massachusetts 01002 (e-mail: jdonahoe@psych.umass.edu).

molecular view, responding during a two-component concurrent schedule is a mixture of at least four operants: responding to Alternative A, responding to Alternative B, responding to A immediately after switching from B, and responding to B immediately after switching from A. Averaging over these different operants yields the overall rate of responding to an alternative. As will become apparent, not distinguishing between these different operants presents difficulties for the GA model. The NN and, less clearly, the CA models have as their target the occurrence of individual responses of single organisms.

The Formal System

The CA model. The CA model differs from most theories of behavior in that its rules (laws) are not the result of prior experimental analyses of fundamental biobehavioral processes but of formal considerations alone. CA models are certainly not alone in being influenced by formal considerations. For example, Baum stated, “when I started studying choice, I didn’t know where it would lead. I was fascinated by the quantitative possibilities. Only later did I realize it suggested a whole new view of behavior” (Baum, personal communication, 2012). In the CA model, an initial linear array of black or white cells is operated on by a set of rules at successive time intervals (time-steps). The rules arise from formal considerations and are not the fruits of prior experimental analysis. The rule that applies to a cell at a time-step is determined by the binary state of that cell (i.e., black or white) and the states of its neighboring cells. This process is allowed to run over a number of time steps with the state of the cell (e.g., black) indicating whether a simulated response occurs at that time-step. Under some sets of rules, the simulated distribution of responses in time resembles the distribution observed in

real organisms (McDowell & Popa, 2009).

The potential correspondence between behavior and the output of a CA is taken to suggest that a theory “without nature in it” may nevertheless have value in behavior analysis. In support of this conjecture, McDowell suggests parallels between the CA model and certain alternative accounts in quantum mechanics that predict the same observations but that differ in their formulations, neither of which includes independently measurable variables. I confess that I am inherently skeptical of appeals to similarities between behavior analysis and quantum mechanics. I recall an article written by a psychotherapist in which sudden changes in a client’s behavior were likened to quantum changes in the energy levels of electrons, with “love” serving as the counterpart of energy. (I cannot retrieve the exact reference, probably a repressed memory!) McDowell’s conjecture is certainly not of this crude type, but the observational bases of quantum mechanics and behavior analysis are so different that they weaken the analogy. The deservedly high status of quantum mechanics rests on its ability to describe and predict an enormous range of observations despite Heisenberg uncertainty and its attendant observational limitations. Accepting the lack of “nature” in quantum mechanics as a guideline for behavioral theory risks mistaking a specific necessity for a general virtue. All of the immediate antecedents of behavior are observable in principle. Whatever value CA models hold for behavior theory remains unclear.

Because the rules of the CA model are not the product of experimental analysis, the CA model does not qualify as a scientific interpretation in Skinner’s sense of the term. Interpretation is sanctioned in behavior analysis when the conditions under which the behavior occurs do not permit the control and measurement

TABLE 1

The three components of a selection process, illustrated by examples from a genetic algorithm model and a neural network model

Components of selection process	Genetic algorithm model	Neural network model
Variable factor	Genetic mutation	Neural thresholds
Selecting factor	Differential survival	Differential reinforcement
Enduring factor	Genes	Synaptic efficacies

of all of the variables required by an experimental analysis (e.g., Skinner, 1957). In such circumstances, the behavior is interpreted ("explained") if the extant circumstances and history of selection can reasonably be thought to include events sufficient to engage fundamental processes previously identified through an experimental analysis. If the net effect of these processes is sufficient to produce the observation as their consequence, then the phenomenon is explained.

The GA model. In the GA model, a response to an alternative is represented by a set of cells, each of which contains a series of bits that constitute a binary number (McDowell, 2004). In the simulations, a reinforcer becomes available according to a two-alternative random-interval schedule after a sufficient number of time-steps have elapsed. The two responses are represented by independent subsets of the set of cells. The GA runs through successive time-steps until responding has stabilized. After a number of simulated organisms have been so trained, pairs of the resulting population of cells are mated in proportion to their obtained number of reinforcers. Variation is introduced into the mating process by simulating genetic crossover and mutation. Thus the cells of the offspring are highly similar, but not identical, to the cells of the parents. Note that the GA model seeks to account for molar relations, but it does so through the instantiation of a molecular process (individual simulated response–reinforcer contiguities).

Although the GA model does not implement a *behavioral* interpretation, it does qualify as an instance of selectionism. In this respect, the GA model shares an explanatory strategy with NN accounts of reinforcement (Donahoe, Burgos, & Palmer, 1993). The observed behavior is the net effect of its history of selection. Selection processes have three components: A population of variants is subject to a selecting factor that favors some variants over others, and the favored variants endure long enough to be subjected to later selection (Donahoe, 2003). As a simulation of conditioning, the GA model is inconsistent with genetics, as McDowell acknowledges; the model is Lamarckian in that the conditioning that occurs in one generation is passed to the next generation of simulated organisms (cf. Baldwin, 1896). As illustrated in Table 1, the genetic components of the GA model have their counterparts in neural networks. Indeed, a simulation of the learning process would more appropriately employ a learning algorithm than a genetic algorithm.

Prediction

After listening to McDowell's presentation at the conference, I read most of his published work on the simulation of performance on schedules of reinforcement using the GA model. As accurately indicated in the article in this volume, the GA model has been implemented in a very wide variety of concurrent ("choice") procedures and has demonstrated an

impressive correspondence between simulated and observed responding. Close correspondence has been found not only with respect to the relative rates of responding but also with respect to the parameters of the best fitting functions and to more detailed aspects of performance. Moreover, the findings are robust, holding over changes in a number of particulars of the simulations, such as the degree of variation (mutation rate) and the initial probabilities of the two responses (the number of cells in the subsets that correspond to the alternatives responses). The GA model convincingly demonstrates that molar measures of responding can arise as the cumulative effect of moment-to-moment contiguities between simulated responses and reinforcers. This finding is consistent with considerable experimental work (e.g., Hinson & Staddon, 1983; Silberberg, Hamilton, Ziriax, & Casey, 1978; cf. Donahoe, 2012).

Be that as it may, the simulation of performance on concurrent procedures with the GA model differs in important respects from Skinner's conjectures about such schedules and differs from some experimental findings as well. As noted earlier, Skinner argued against accepting responding to an alternative as a measure of the strength of a single operant. Instead, such a measure conflated the strength of two operants: responding to an alternative and switching to that alternative. In the GA simulations, only two operants are considered (responding to one or the other alternative). Switching operants has no representation in the simulations. Instead, switching occurs when the computer program probabilistically selects the other alternative on the basis of its momentary strength. In short, switching behavior is not acquired, it is "innate." In Bertrand Russell's happy phrase, incorporating switching in the CA model by postulating its existence "has many advantages; they

are the same as the advantages of theft over honest toil" (Russell, 1919, p. 71). Experimental analysis indicates that switching is not innate but requires the explicit reinforcement of switching operants if concurrent schedules are to produce matching. When two operants of differing strengths are acquired without including a history of reinforcement for switching, organisms evince an absolute preference for the stronger operant (Crowley & Donahoe, 2004). Matching does not occur. However, in the GA model, if two operants were independently trained and then made concurrently available, responding would still show matching because an alternative is chosen in proportion to its relative strength. This is inconsistent with experimental evidence.

To confront this difficulty, it might be supposed that the CA model could simply be extended to include four operants, responding to each of the two alternatives and switching between them. Experimental work has shown that these four operants may be acquired on independent reinforcement schedules (MacDonall, 2009). However, such a simulation by the CA model would encounter a new difficulty: What would be the reinforcer for switching responses? Except on those very rare occasions when the switching response was immediately followed by a reinforcer, switching would not be reinforced. And, if a changeover delay were simulated, which is a ubiquitous characteristic of experimental work on concurrent schedules, switching would never be reinforced. In experimental realizations of concurrent schedules, switching behavior is maintained by conditioned reinforcers provided by the appearance of stimuli from the switched-to alternative (Findley, 1958; McDevitt & Williams, 2010). The GA model has no inherent mechanism for simulating conditioned reinforcement. In chained as well as concurrent schedules, the

stimulus of a component serves as a conditioned reinforcer for responding to that component. In previous work on chained schedules, the GA model was modified so that the stimulus of a component was defined as a reinforcer for switching (McDowell, Soto, Dallery, & Kulubekova, 2006). As with switching behavior in concurrent schedules, this effect was imposed externally in the simulation and does not emerge from the GA model itself. Russell's words again come to mind. Note also that introducing conditioned reinforcement into the GA model on the basis of behavioral evidence places "nature" squarely in the model.

Simulations using the GA model do not meet Skinner's definition of a scientific interpretation, but the model does convincingly demonstrate that molar measures of responding may arise as the cumulative effect of momentary relations between individual responses and reinforcers. Molar response measures and the outcome of simulations using the GA model are in exquisite agreement. On these grounds alone, the GA model makes an important contribution to our understanding of concurrent schedules: The momentary processes simulated in the GA model are sufficient to produce molar matching even though they do not simulate the momentary processes found in individual organisms.

Explanation

As noted by Killeen (2013), scientific explanation is achieved by what the philosopher of science Carl Hempel (Hempel & Oppenheim, 1948) called a *covering law*. That is, an observation is explained if it is an implication of premises that include a natural law (i.e., a law "having nature in it"). Clearly, the GA model does not provide an explanation in this sense, nor does it claim to. Instead, the validity of the GA model rests solely on its undeniable ability

to produce outcomes that agree in major respects with experimental observations of molar measures of performance on concurrent schedules.

The attempt to explain behavior without reference to the biobehavioral processes of which the outcome is a product is reminiscent of earlier attempts to understand evolution before its genetic mechanisms were uncovered. In the years immediately preceding the rediscovery of Mendel's work, biologists such as Weldon and statisticians such as Pearson favored a purely mathematical approach to natural selection.

A "theory of mechanism" was no more necessary than in the case of selective mortality. A purely statistical approach to heredity, based on correlation, would suffice. ... The direction and rate of evolution could be determined without introducing any theory of the physiological functions of the organisms investigated. (Weldon, 1895, p. 397, cited in Gayon, 1998, p. 214)

"Science is a conceptual *description* and classification, ... a theory of symbols which economizes thought. It is not an explanation of anything" (Pearson, 1900, p. 205).

The GA model stands in contrast to simulations using NN models that, to varying degrees, incorporate observations from relevant behavioral and neural sciences. NN models are criticized because they do not incorporate all such observations. This criticism makes the perfect enemy of the good and is misplaced, as the following statement indicates.

It is unlikely—and even undesirable—for any simulation to instantiate *all* biobehavioral knowledge because the simulation would then be as complex (and as difficult to understand) as the processes whose outcomes it simulates. Instead, it is sufficient for the simulation to capture the phenomena to the desired particularity, employing those processes and structures minimally necessary to accomplish that goal, while not violating any known experimental finding. To the extent that the outcome of the simulation is consistent with the phenomenon, then—to that extent—are the simulated processes competent to produce the phenomena.

In short, the simulation demonstrates that those processes "explain" the phenomena. Of course, as the range of phenomena encompassed by a simulation increases, the need to incorporate additional findings may also increase. (Donahoe, 1997, p. 351)

This view of explanation in biobehavioral science is consistent with simulations in population genetics in which not all potentially relevant environmental and genetic factors inform every simulation. Moreover, it comports with Skinner's very early injunctions about the appropriate level of analysis at which to explain instances of behavior.

As we proceed with the gradual restriction of a preparation, noting a corresponding improvement in the consistency of our data, the point at which an adequate consistency is reached does not coincide with the final complete restriction of all properties of the preparation. ... This is a practical rule, which does not pretend to go beyond the limits of our present degree of precision. ... It would be idle to consider the possibility of details which have at present no ... importance. ... We are here interested only in the degree of consistency which can be obtained while they are still by no means completely determined. (Skinner, 1935, p. 516)

MODELS IN BEHAVIOR ANALYSIS

The present discussion is guided by Skinner's judgment that behavior analysis is "a rigorous, extensive, and rapidly advancing branch of *biology* [emphasis added]" (Skinner, 1974, p. 231). (I acknowledge that some who regard their views as falling under the behavior-analytic tent, such as interbehaviorists, disagree with this characterization of the field; Kantor, 1958; Reese, 1996.) As a branch of biology, behavior analysis is no less independent of, and no less interdependent on, such sister sciences as neuroscience and biochemistry. Theories in behavior analysis should be constrained and informed by empirical findings from its own efforts but not be inconsistent with allied biological disciplines. Shortly

before his death, Margaret Vaughan asked Skinner to speculate about future areas of research that were important for behavior analysis (Vaughan, personal communication, 1994). Among those Skinner identified was the neural basis of conditioning.

Strategies of Model Evaluation

Killeen's discussion characterizes the goal of theory as facilitating mutual communication between the model and the data. That is, given the model, present data should be implied (*deduction*); given the data, the theory should be consistent (*induction*) or productively modified (*abduction*); and, if the theory is fruitful, given the model, future observations should be implied (*emergence*). The GA model differs in important respects from this view of the relation among theory, model, and data. However, the model does generate many of the effects of concurrent schedules when simulating selection by reinforcement.

As a measure of mutual communication between model and data, Killeen suggests Shannon's information measure. For example, given an observation, what information is conveyed about the "truth" of the model? The observation that has the highest probability of changing the truth value of the model (whether increasing or decreasing it) is the most informative. This example falls within the province of Bayes's inverse probability principle: The a posteriori probability of a model given the observation (i.e., the probability of the model after the observation) may be compared with its a priori probability. The greater the difference in these probabilities, the greater the information value of the observation with respect to the model. Information theory provides a measure of this change in probability.

A major thrust of the information metric is captured in an early paper by the geologist Thomas Chamberlin

(1890/1965; see also Platt, 1964). Chamberlain favored experiments whose observations were sufficient to decide among competing models. He called this strategy the method of *multiple working hypotheses*; Platt referred to the strategy as the method of *strong inference*. If there are competing models and the experiment is well chosen, then the information about the models conveyed by the experiment is maximized. For example, if there are only two competing models, then the maximum information value transmittable by an experiment is one bit. If there are four competing models, then two bits are transmittable by an observation that decides among them, and so on. Of course, the situation in the real world is often not that clear-cut: The list of hypotheses (models) may not be exhaustive, and the data may be noisy. Nonetheless, the idealized circumstances envisioned in Killeen's essay sensitize us to some of the important issues that should guide research.

Fundamental Shortcoming of "Natureless" Theories

The most serious deficiency of theories that do not have "nature" in them is that they do not exploit experimental findings that arise from that science or related sciences. Each model is a formal construction unto itself. Consider the GA model. What are its implications for phenomena beyond laboratory observations of behavior under concurrent random-interval schedules? Such schedules have been the object of considerable experimental study within behavior analysis, but what are the model's implications for other phenomena in the laboratory (e.g., stimulus control) and in that greater world beyond the laboratory? As but one problematic example, blocking is arguably the most important discovery in conditioning during the past 40 years (Kamin, 1968, 1969; see Papini &

Bitterman, 1990). The blocking phenomenon has deepened our understanding of the nature of the critical reinforcing event: It is the contiguity of the response with the *behavioral change* evoked by the reinforcer that is critical, not the occurrence of the reinforcing stimulus per se (Donahoe & Vegas, 2004). However, if the experimental design that produces blocking were simulated with the GA model, conditioning could either not be simulated because stimuli have no representation within the model or, if representations of stimuli were introduced, then conditioning of responses to the newly introduced stimulus would occur because temporal contiguity of a selected response with the reinforcer is sufficient to strengthen the response.

Although interpretations of schedules of reinforcement have undeniable value, it is well to recall the purpose for which schedules were devised: "Under a given schedule of reinforcement, it can be shown that at the moment of reinforcement a given set of stimuli will usually prevail. A schedule is simply a convenient way of arranging this" (Ferster & Skinner, 1957, p. 3). Given the technology of the time, the classic schedules achieved this goal as well as relays, timers, and stepping switches allowed. Nevertheless, Herb Jenkins's admonition on the subject of schedules is pertinent: "Schedules may be viewed as contrivances to be used in order to improve our understanding of conditioning principles. ... They are an invention and it is possible to choose whether or not to analyze the effects they produce" (Jenkins, 1970, p. 107).

The limitations of "natureless" theory are more apparent and arguably more important when interpreting behavior in the world outside the laboratory. As one example, conditioned reinforcement is an emergent property of the NN model (Donahoe, 1997), as anticipated by behavioral research (Dinsmoor, 1950; Keller &

Schoenfeld, 1950). Because of the intimate correspondence between speech sounds and the articulatory responses that produce these sounds, behavior analysis and the NN model provide a mechanism whereby language acquisition is not dependent on immediate external reinforcers: If the prior history has established speech sounds as discriminative stimuli for nonverbal behavior, then vocalizations are automatically reinforced to the extent that they produce auditory stimuli that approximate those speech sounds (Donahoe & Palmer, 1994/2009, pp. 318–319, Palmer, 1996, Vaughan & Michael, 1982). The GA model is silent on such matters. Moreover, the virtues of the GA model are not necessarily beyond the reach of an NN model. Calvin (2012), working in McDowell's laboratory, demonstrated through simulation research that "the Donahoe neural networks generated behavior similar to that of biological organisms" (p. 22) with the exponentiated hyperbola accounting for 97% of the variance.¹

Finally, when dealing with networks that have "nature in them," parameter values can potentially be determined through independent experimental analysis, and are not

simply estimated according to goodness-of-fit criteria.

Toward Comprehensive Biobehavioral Simulations

A comprehensive interpretation of behavior requires an integration of behavioral research with research at many levels of analysis. Behavioral principles and findings stand on their own merits, but they enrich, and are enriched by, their relation to other biological sciences and to the behavioral findings of cognitive psychology.

Behavior analysis and biology. Figure 1 depicts the interplay between behavior analysis and its related biological sciences (Donahoe, 1997). Because of the interdependence of these sciences, discussion of their dynamical relations could begin at any of several points. Here, we begin with the organism as represented by the network architecture. The network architecture is the structure of the neural network and consists of its inputs, outputs, and interior units with their interconnections. The architecture is the product of natural selection simulated by a connections between units and is informed by both behavioral and neural research (Donahoe, 1997; Donahoe et al., 1993; Donahoe & Vegas, 2004). The environmental algorithm implements the contingencies between environmental events, particularly the reinforcement function (the schedule of reinforcement), the shaping function (the requirements for a reinforced response), and the fading function (the effect of behavior on the subsequent environment). As a result of the simulated history of reinforcement, the strengths of connections between units are modified over time to produce a trained network. The trained network constitutes one member of a population of trained networks that differ in their response to the contingencies of reinforcement. The fitness function determines the probability that the simulated genes

¹ Units in neural networks of the type we have developed (i.e., selection networks; Donahoe et al., 1993) are activated according to a logistic function. The logistic probability density function has a "fat tail" (i.e., more frequent high values than the normal distribution) that mimics the firing distribution of neurons. Output units of the network can be viewed as "summing" the activations of the units with which they are connected. In accordance with the generalized central limit theorem, the sum of a set of such functions is a power function (exponentiated hyperbola), which is consistent with the frequent occurrence of power functions in behavioral data (cf. Stumpf & Porter, 2012). The proportion of variance accounted for by the NN model in Calvin's simulations would likely have been even higher if the network had more units; that is, had the sum been over a larger number of units.

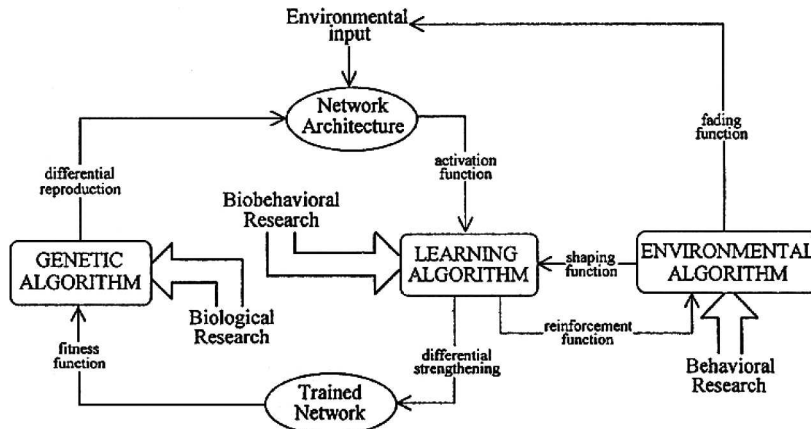


Figure 1. Toward a comprehensive simulation of the relation between behavior analysis and the allied sciences of neuroscience and genetics.

and developmental processes that produced the architecture contribute to the next generation of networks. The genes contribute to the next generation in proportion to the effectiveness of the behavior mediated by the networks whose construction they guided (differential reproduction). With this system, which is informed by behavior analysis, neuroscience, and genetics, simulations have produced neural networks that are effective after training with the reinforcement schedule. Using this schema, some of the effects of Pavlovian and operant contingencies have successfully simulated phenomena such as acquisition, extinction, delay of reinforcement, revaluation, reacquisition, and Kamin blocking, among others (e.g., Donahoe & Burgos, 2005; Donahoe et al., 1993). The system is informed by the fruits of experimental analysis. “Nature” is in it, but not all potentially relevant nature.

Behavior analysis and complex behavior. The present commentary has focused on the characteristics of GA and NN simulations of behavior because of the nature of the target articles. However, NN simulations are not the only approach that meets Skinner’s criteria for scientific interpretation. Three complementary strate-

gies of interpretation are distinguishable (Donahoe & Palmer, 1994/2009): The first is *verbal interpretation*, in which ordinary language specifies the basic processes that might produce the complex behavior. Skinner’s *Verbal Behavior* (1957) is a notable example of this strategy. The second strategy is *organismic interpretation*, whereby an organism that does not display the complex behavior is exposed to a history of reinforcement that verbal interpretation suggests is sufficient to produce the behavior. Epstein and Skinner’s work (1982) on “insight” in the pigeon exemplifies this strategy. Finally, there is *formal interpretation*, which uses the special verbal behavior of mathematics and logic (often implemented with a digital computer) to simulate the complex behavior. NN simulations fall within this strategy but so do simulations that implement solely behavioral principles. Hutchison’s (2012) studies simulating the cumulative effects of multiple discriminated operants on verbal behavior and on motor control provide examples of this strategy (cf. Epstein, 1985, 1991). Within cognitive psychology, work on parallel distributed processing is conceptually most closely related to formal interpretation in behavior analysis (e.g., Elman, 1995; McClelland, 2012). However,

the elementary processes from which the complex behavior arises are generally not those identified by behavior analysis and are sometimes inconsistent with neuroscience (Donahoe & Palmer, 1989). Although NN simulations informed by neuroscience are best suited to bridge the gap between behavior analysis and the biological sciences, operant-level simulations are perhaps better positioned to explore the implications of behavioral principles for the complex behavior studied in cognitive (i.e., inferred process) psychology. Together, these diverse interpretive strategies hold the as-yet-unrealized promise of a comprehensive, science-based understanding of complex behavior, including human behavior.

REFERENCES

- Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30, 441–451.
- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137–153.
- Calvin, N. (2012). *Accordance of the unified reinforcement's model of behavior with the modern quantitative law of effect*. Unpublished master's thesis, Emory University, Atlanta, GA.
- Chamberlain, T. C. (1965). The method of multiple working hypotheses. *Science*, 148, 754–759. (Original work published 1890)
- Crowley, M. A., & Donahoe, J. W. (2004). Matching: Its acquisition and generalization. *Journal of the Experimental Analysis of Behavior*, 82, 143–149.
- Dinsmoor, J. A. (1950). A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *Journal of Experimental Psychology*, 40, 458–472.
- Donahoe, J. W. (1997). The necessity of neural networks. In J. W. Donahoe & V. P. Dorsel (Eds.), *Neural-network models of cognition*. New York, NY: Elsevier.
- Donahoe, J. W. (2003). Selectionism. In K. A. Lattal & P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 103–128). Dordrecht, Netherlands: Kluwer Academic.
- Donahoe, J. W. (2012). Origins of the molar-molecular divide. *European Journal of Behavior Analysis*, 19, 196–200.
- Donahoe, J. W., & Burgos, J. E. (2005). Selectionism: Complex outcomes from simple processes. *Behavioural Brain Research*, 28, 429–430.
- Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior*, 60, 17–40.
- Donahoe, J. W., & Palmer, D. C. (1989). The interpretation of complex human behavior. *Journal of the Experimental Analysis of Behavior*, 51, 399–416.
- Donahoe, J. W., & Palmer, D. C. (2009). *Learning and complex behavior*. Richmond, MA: Ledge-top Publishers. (Original work published 1994) Retrieved from <http://lcb-online.org/>
- Donahoe, J. W., & Vegas, R. C. (2004). Pavlovian conditioning: The CS-UR relation. *Journal of the Experimental Psychology: Animal Behavior Processes*, 30, 17–33.
- Elman, J. L. (1995). Language as a dynamical system. In R. Port & T. van Gelder (Eds.), *Mind as motion* (pp. 195–224). Cambridge, MA: MIT Press.
- Epstein, R. (1985). Animal cognition as the praxis views it. *Neuroscience & Biobehavioral Reviews*, 9, 623–630.
- Epstein, R. (1991). Skinner, creativity, and the problem of spontaneous behavior. *Psychological Science*, 2, 362–370.
- Epstein, R., & Skinner, B. F. (1982). *Cognition, creativity, and behavior*. Champaign, IL: Research Press.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York, NY: Appleton-Century-Crofts.
- Findley, J. D. (1958). Preference and switching under concurrent scheduling. *Journal of the Experimental Analysis of Behavior*, 1, 123–144.
- Gayon, J. (1992). *Darwin's struggle for survival* (M. Cobb, Trans.). Cambridge, UK: Cambridge University Press.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135–175.
- Hinson, J. M., & Staddon, J. E. R. (1983). Hill-climbing by pigeons. *Journal of the Experimental Analysis of Behavior*, 39, 25–47.
- Hutchison, W. R. (2012). The central role for behavior analysis in modern robotics, and vice versa. *The Behavior Analyst*, 35, 29–35.
- Jenkins, H. M. (1970). Sequential organization in schedules of reinforcement. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 63–109). New York, NY: Appleton-Century-Crofts.
- Kamin, L. J. (1968). Attention-like processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior* (pp. 9–31). Miami, FL: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.
- Kantor, J. R. (1958). *Interbehavioral psychology*. Bloomington, IN: Principia Press.

- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology: A systematic text in the science of behavior*. East Norwalk, CT: Appleton-Century-Crofts.
- Killeen, P. (2013). The structure of scientific evolution. *The Behavior Analyst*, 36, 325–344.
- MacDonall, J. S. (2009). The stay/switch model of concurrent choice. *Journal of the Experimental Analysis of Behavior*, 91, 21–39.
- McClelland, J. L. (2012). Emergence in cognitive science. *Topics in Cognitive Science*, 2, 751–770.
- McDevitt, M. A., & Williams, B. A. (2010). Dual effects on choice of conditioned reinforcement frequency and conditioned reinforcement value. *Journal of the Experimental Analysis of Behavior*, 93, 147–155.
- McDowell, J. J. (2004). A computational model of selection by consequences. *Journal of the Experimental Analysis of Behavior*, 81, 297–317.
- McDowell, J. J. (2013). Representations of complexity: How nature appears in our theories. *The Behavior Analyst*, 36, 345–359.
- McDowell, J. J., & Popa, A. (2009). Beyond continuous mathematics and traditional scientific analysis: Understanding and mining Wolfram's *A New Kind of Science*. *Behavioural Processes*, 81, 343–352.
- McDowell, J. J., Soto, P. L., Dallery, J., & Kulubekova, S. (2006). A computational theory of adaptive behavior based on an evolutionary reinforcement mechanism. In M. Keijzer (Ed.), *Proceedings of the 2006 Conference on Genetic and Evolutionary Computation* (pp. 175–182). New York, NY: ACM Press.
- Morse, W. H. (1966). Intermittent reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 52–108). Englewood Cliffs, NJ: Prentice Hall.
- Palmer, D. C. (1996). Achieving parity: The role of automatic reinforcement. *Journal of the Experimental Analysis of Behavior*, 65, 289–290.
- Papini, M. R., & Bitterman, M. E. (1990). The role of contingency in classical conditioning. *Psychological Review*, 97, 396–403.
- Pearson, K. (1900). *The grammar of science* (2nd ed.). London, UK: A and C Black.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347–353.
- Reese, H. W. (1996). How is physiology relevant to behavior analysis? *The Behavior Analyst*, 19, 61–70.
- Russell, B. (1919). *Introduction to mathematical philosophy*. London, UK: Allen and Unwin.
- Silberberg, A., Hamilton, B., Ziriaux, J. M., & Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes*, 4, 368–398.
- Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *Journal of General Psychology*, 12, 40–65.
- Skinner, B. F. (1957). *Verbal behavior*. New York, NY: Appleton-Century-Crofts.
- Skinner, B. F. (1974). *About behaviorism*. New York, NY: Random House.
- Stumpf, M. P. H., & Porter, M. A. (2012). Critical truths about power laws. *Science*, 335, 665–666.
- Vaughan, M., & Michael, J. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism*, 10, 217–227.